cross forest

Co-financed by the European Union
Connecting Europe Facility

User Specifications and Requirements

Deliverable D3.1: Public

Keywords: Forest fire, Wood quality, modellization, simulation

CROSS Harmonization & HPC modelization of FOREST Datasets

Action number: 2017-EU-IA-0140

Agreement INEA/CEF/ICT/A2017/1566738

Table of Contents

List of Figures

List of Tables

Contractual Date of Delivery to the EC:	May 2019
Actual Date of Delivery to the EC:	31/05/2019
Editor(s):	Felipe Bravo (Uva) Cristóbal Ordóñez (Uva) Asuncion Roldán (TRAGSA) Jesús Estrada (TRAGSA) Victor González (TRAGSA) Alvaro Carrillo (TRAGSA)
Contributor(s):	

DOCUMENT HISTORY

Version	Version date	Responsible	Description
0.1	29/05/2019	UVa	Draft for Uva TRAGSA comments
0.2	30/05/2019	UVa	Draft for Cross-Forest consortium comments

1 Introduction

This document defines the procedures developed in the Pilots and practical scenarios of Cross-Forest, Frame and Cambric, aimed to design and develop the practical use cases for implementation of HPC services applied to forestry environment.

For each pilot, information about the available data sources will be provided, the sequence of steps to be carried out, the pretended outputs and the necessary transformations of the data sources.

The utilization of cross-border forest datasets in these two scenarios represents a fundamental progress in cross-border information sharing in Europe, allowing better forest management and protection against fires and other hazards.

The proposed pilots will use datasets gathered by the Directorate General for Territory – DGT (Portugal) and the Agriculture and Environment Ministry – MAPA (Spain) to create harmonized data models oriented towards common publication of forest information.

In this sense, it is essential to define what processes are involved and who should participate. In order to achieve the main objective of the project, we must capture the sequence of actions performed or to be carried out by an external entity on the system to achieve a quantifiable goal.

1.1 Purpose

The main purpose of this document is to show how pilots interact with the endpoint, in the Cross-Forest Project scenarios. To ensure that the endpoint repositories, external data sources and sequences can answer the project team questions regarding forest management and fire spreading scenarios.

Both pilots, Frame and Cambric, aim to use already developed simulators FRAME and Simanfor, but adapted to an HPC environment so as to enhance their capabilities.

Both pilots also aim to adapt their input and output databases to Linked Open Data, as far as possible.

1.2 Scope

In this first stage, we meet SCAYLE team in order to discuss different options to be used in Calendula/HPC environment. The scope of this document aims:

a) To guide the project development, through the identification of information required and how to make the necessary transformations in each pilot.

b) To validate the project development, through the pilots, identifying the required sources, formats and expected outputs.

c) To identify sources (internal and external data) that allows obtaining the expected data in each test.

d) To guarantee the access to internal and external sources and data, that will allow obtaining the expected data.

1.3 Definitions, acronyms and abbreviations

AEMET	Spanish Meteorological Agency
CALENDULA	Supercomputer located at SCAYLE
CAMBRIC	CAlidad de la Madera en Bosques mIxtos (Wood quality in mixed forest)
CBD	Crown Bulk Density
CEF	Connecting Europe Facility
CFIS	Crown fire initiation and spread
CORDIS	Community Research and Development Information Service
DGT	Direção-Geral do Território
DTM	Digital Terrain Model
EC	European Commission
EMERCARTO	
EU	European Union
FRAME	Forest fiRes Advanced ModElization
GIS	Geographic information system
HPC	High performance Computing
ICT	Information and communications technology
ICNF	Instituto de Conservação da Natureza e Florestas
IGN	National Geographic Institute (Spain)
INEA	Innovation and Networks Executive Agency
IPMA	Portuguese Institute for the Sea and Atmosphere
LiDAR	Laser Imaging, Detection And Ranging
LOD	Linked Open Data
MAPA	Agriculture and Environment Ministry (Spain)
NFI	National Forest Inventory
SCAYLE	Fundación Centro de Supercomputación Castilla y León (Castilla y León Supercomputing Center)
SIMANFOR	Support system for simulating Sustainable Forest Management Alternatives
TRAGSA	Empresa de Transformación Agraria, S.A.
UCO	
UVA (IuFOR)	University of Valladolid (Institute for Research in Sustainable Forest Management)
UVA (GSIC/EMIC)	University of Valladolid (Group of Intelligent & Cooperative Systems / Education, Media, Informatics & Culture)

2 Pilots specifications and requirements

Here we will describe separately the different steps followed by each Pilot Project.

For both pilots, the tools to be applied consist in data models of the Forest Inventories and Forest maps available for each country (Portugal and Spain). These data sources will be processed to obtain the expected data in each pilot.

Figure 1 - Cross-Forest logo

2.1 CAMBRIC: CAlidad de la Madera en Bosques mIxtos

2.1.1 Overall Perspective

The aim of CAMBRIC is to forecast wood quality in mixed forest at big surface levels (over municipality). Using Big Data as source information provided by National Forest Inventory, Remote Sensing or LiDAR, a set of Yield Models adequate for each stand, and different management scenarios, we will develop a dynamic information platform with wood production for different end uses.

There will be four out of the five silvicultural scenarios proposed by Duncker et Al. (2012): (1) Passive: Unmanaged forest nature reserve; (2) Low: Close-to nature forestry, (3) Medium: Combined objective forestry, (4) High: Intensive even-aged forestry and (5) Intensive: Short-rotation forestry.

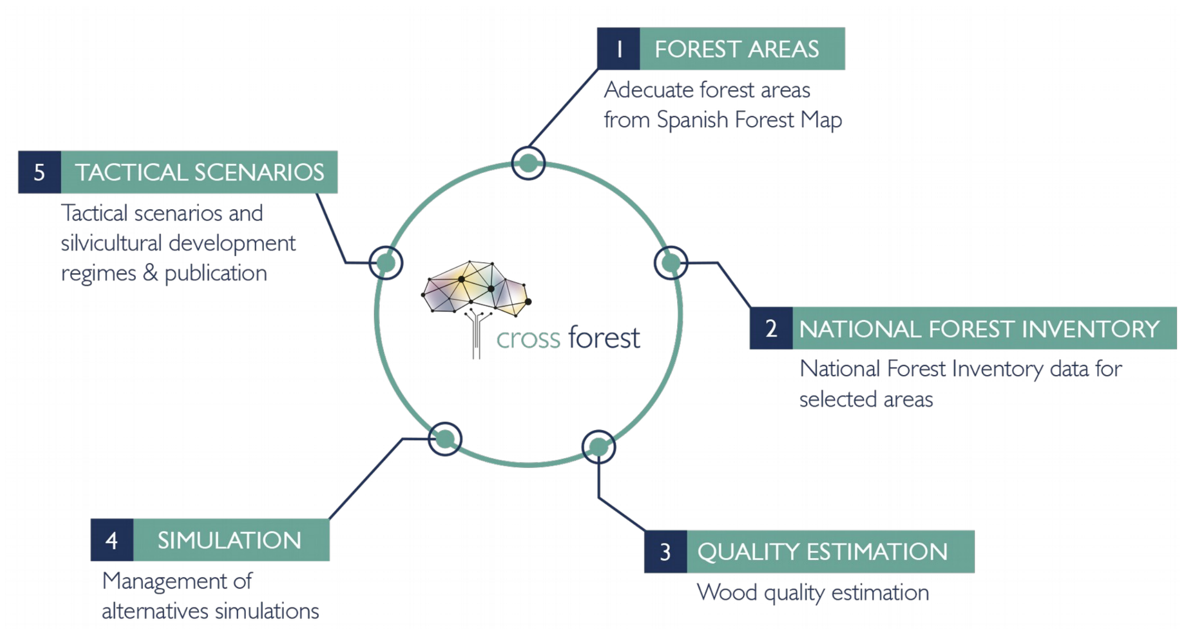

Figure 2 - CAMBRIC pilot overall perspective

These management schemes will be simulated across a gradient of forest mixtures (from pure to multi-species stands) for the target species. Simulations will allow to classify timber products on a log size and stem knot-free length basis. The work-flow (see Figure 2) will be (i) select adequate areas with the information provided by the Spanish Forest Map covers, (ii) extract forest inventory data for the selected areas, (iii) estimate wood quality by applying stem taper equations to inventory data, (iv) simulate different silvicultural alternative for the Spanish NFI data to explore the impact of species mixture on wood quality and (v) develop tactical scenarios for the different silvicultural regimes and publish output data following the Linked and Open data approach defined in the project.

The following taxa will be selected as model species to analyze the availability of wood and its quality for different end uses: Pinus sylvestris, Pinus pinaster, Pinus nigra, Fagus sylvatica, Quercus pyrenaica, Quercus robur and Quecus petraea, and Pure and mixed stands of these species will be analysed.

2.1.2 Groups by species composition.

First we should select available growth and yield models for each taxa. Table 1 describes the different species and already fitted relationships.

Table 1 – Species composition and models available

id	Species composition	Model (Bravo et al, 2012)						
		Site index	Growth	Mortality	Ingrowth	Taper	Volume biomass	Auxiliary
1	Pinus sylvestris (Psy) pure	SI	d.g, h.g	Mt	Ig	t	V, B	bt, h-d, crw
2	Pinus pinaster (Ptr) pure	SI	d.g, h.g	Mt	Ig	t	V, B	bt, h-d, crw
3	Pinus pinaster atlantica (Ptr-A)	Galicia	d.g, h.g	-	-	t	V, B	bt, h-d, crw
4	Pinus nigra (Png) pure	SI	-	-	-	C y L	C y L	bt
5	Fagus sylvatica (Fsy) pure	SI	-	-	-	C y L / Navarra	C y L / Navarra	bt
6	Quercus pyrenaica (Qpy) pure	SI	-	-	-	C y L	C y L	bt

7	Quercus robur (Qrb) pure	Galicia	Galicia	Galicia	-	Galicia	Galicia	bt, h-d
8	Quercus petrea (Qpt) pure	-	-	-	-	-	-	bt
9	Psy + Ptr	-	-	-	-	-	-	-
10	Psy + Png	-	-	-	-	-	-	-
11	Psy + Fsy	-	-	-	-	-	-	-
12	Psy + Qpy	-	-	-	-	-	-	-
13	Psy + Qrb	-	-	-	-	-	-	-
14	Psy + Qpt	-	-	-	-	-	-	-
15	Ptr + Png	-	-	-	-	-	-	-
16	Ptr + Qpy	-	-	-	-	-	-	-
17	Png + Qpy	-	-	-	-	-	-	-
18	Fsy + Qrb	-	-	-	-	-	-	-
19	Fsy + Qpt	-	-	-	-	-	-	-
20	Qrb + Qpt	-	-	-	-	-	-	-
21	Qrb + Qpy	-	-	-	-	-	-	-
22	Qpt + Qpy	-	-	-	-	-	-	-
23	Others	-	-	-	-	-	-	-

Equations reference: Site index equations (SI), Diameter and basal area growth functions (d.g), Height growth functions (h.g), Volume growth functions (h.g), Mortality models (Mt), Ingrowth models (Ig), (V) Volume and biomass (B) equations, Bark thickness or bark percentage (bt), Height-diameter relationships (h-d), Crown equations (crw)

For the species composition we lack equations, we will develop new ones; For mixed stands we will also develop new models following Riofrío et Al. 2019 developed for Scotch' and Mediterranean' pine mixture, but whether this is not possible we will use the ones provided for mono-specific stands.

2.1.3 Area by species selection and NFI data extraction

Using the shape files obtained through the Spanish Forest Map information, the complete list of patches for each species and mixtures composition will be selected.

With this information, the NFI plots included in this patches will be selected. After checking characteristics of minimum basal area, species correspondence and coherence in data, a first version of input data for simulations will be achieved.

2.1.4 Data input format.

Input data for simulations should be provided in "csv" format to be easily processed at HPC. Data should be stored at tree and stand level in two different files, but with common variables allowing relationships between them.

Table 2 – Stand level input variables

name	detail	name	detail
date	Date of plot setup	D_MAX	Maximum diameter
parcelaIFN	Type of plot	D_MIN	Minimum diameter
superficiePARCELA	Extension of the plot	H_MEDIA	Mean height of trees
expan	Expansion factor	DM_COPA	Mean crown diameter
provincia	Administrative indicator	DG_COPA	Quadratic mean crown diameter

ID_INVENTARIO	Identifier of the inventory	I_REINEKE	Reineke index
ID_PARCELA	Identifier of the plot	I_HART	Hart index
A_BASIMETRICA	Basal area	FCC	Crown coverage
H_DOMINANTE	Dominant height	SI	Site index
N_PIES	Number of trees per plot	Municipio	Administrative division
N_PIESHA	Number of trees per hectare	Monte	Forestry information
EDAD	Age, average of tree age	CoordeX	Absolute coordinate X
D_MEDIO	Mean diameter of trees	CoordeY	Absolute coordinate Y
D_CUADRATICO	Quadratic mean diameter	Altitud	Mean height above sea level
D_DOMINANTE	Dominant diameter	Taxa	Identifier of the species group in this project. This variable allow model selection.

We will use another 10 stand level unnamed variables, VAR_1 to VAR_10, to store input and/or output information not standardize for all models, but necessary for future models.

Table 3 – Tree level input variables

name	detaill	name	detaill
date	Date of plot setup	VCC	Tree volume above bark
provincia	Administrative indicator	VSC	Tree volume under bark
ID_INVENTARIO	Identifier of the inventory	IAVC	Volume annual increment
ID_PARCELA	Identifier of the plot	VLE	Fuel wood volumen
ARBOL	Identifier of the tree	BAL	Basal area of larger trees
NUMEROINDIVIDUOS	Number of trees to represent, for class model. Usually value 1	SECCION_COPA_MAXIMA	Crown projection
ESPECIE	species	EDAD_BASE	Age
DIAMETRO_1	First dbh measured	DIAMETRO_MIN	
DIAMETRO_2	Second dbh measured	DIAMETRO_4	Diameter at 4 m height
CALIDAD		FCV	Crown to height ratio
FORMA		ALTURA_BC	Height to crown base
ALTURA		ALTURA_MAC	Height to maximum crown width
PARAMESP		ALTURA_RM	Height to first dead branch
OBSERVACIONES	observations	ALTURA_VV	Height to first alive branch
DAP	Average of first and second dbh	ALTURA_TOC	Height of the stump
CORTEZA_1	First bark thickness measured	ANCHO_CM_1	crown width 1
CORTEZA_2	Second bark thickness measured	ANCHO_CM_2	crown width 2
CORTEZA	Bark thickness	RADIO_C_1	crown radius 1
CIRCUNFERENCIA	Perimeter at dbh height	RADIO_C_2	crown radius 2
EXPAN	Expansion factor	RADIO_C_3	crown radius 3
ESBELTEZ	Slenderness, ht/dbh, dimensionless	RADIO_C_4	crown radius 4

SEC_NORMAL	Tree basal area	LCW	Maximum crown width
EDAD130	Age at 1.3 m height	Coord_X	Relative x coordinate
CLASE_SOCIOLOGICA	larger than the average trees D=Dominant trees. C-D = Co-Dominant trees. I=Intennediate trees. Op=Oppressed trees. S=Suppressed trees	Coord_Y	Relative y coordinate

We will use another 10 tree level unnamed variables, VAR_1 to VAR_10, to store input and/or output information not standardize for all models, but necessary for future models.

A crucial variable for performing simulations is the age, if not possible at tree level at least at stand level, provided the stands we will use belongs to the same class of age. In order to be able to estimate stand age we will also develop age equations relationships with dominant height and dominant height difference between inventories.

2.1.5 Simulations.

The simulations of alternatives will be performed in CALENDULA Supercomputer. Two tasks should be carried out, adaptation of the simulator and selection of the scenario for silvicultural alternatives.

2.1.6 Simulation environment

The simulations will be performed with a software solution flexible enough to easily adapt different equations for each pure or mixed species stands. Simanfor perfectly fits this requirement, but unfortunately it was developed with incompatible software technology, so it is completely necessary to reprogram a new version compatible with HPC and with usual servers. This task will be externally performed by former developer of the tool, sngular.

2.1.7 Silvicultural alternatives

The following 4 different alternative scenarios will be simulated following methodology by (DUNCKER et al., 2012), depending on taxa characteristics:

1. Passive scenario (Unmanaged Forest Nature Reserve), consisting in lack of management, that is, no cut planned along the life of the stand.
2. Low scenario (Close to nature forestry), consists in carrying out a management imitating the natural processes that occur in the forest, allowing exploitation of wood, although in small quantities, avoiding natural mortality. The operations must be compatible with the ecological functions of the forest.
3. Medium scenario (Combined objective forestry), two objectives are pursued, wood production and the achievement of other economic and ecological benefits.
4. High scenario (intensive even-aged forestry), whose main objective is the production of wood. Only *Pinus pinaster atlantica* taxa lays under the scope of this scenario (see Figure 2).
5. Intensive scenario (Short rotation forestry). This option is not considered because suitable species are out of the scope of this project, i.e. *Eucaliptus* of *Populus* species.

All simulations will start around the age of 30 years and end between 80 and 140 years except the high scenario which starts at 20 years and ends at 40. We will avoid the regeneration period and early stage of the stand due to lack of suitable models and information.

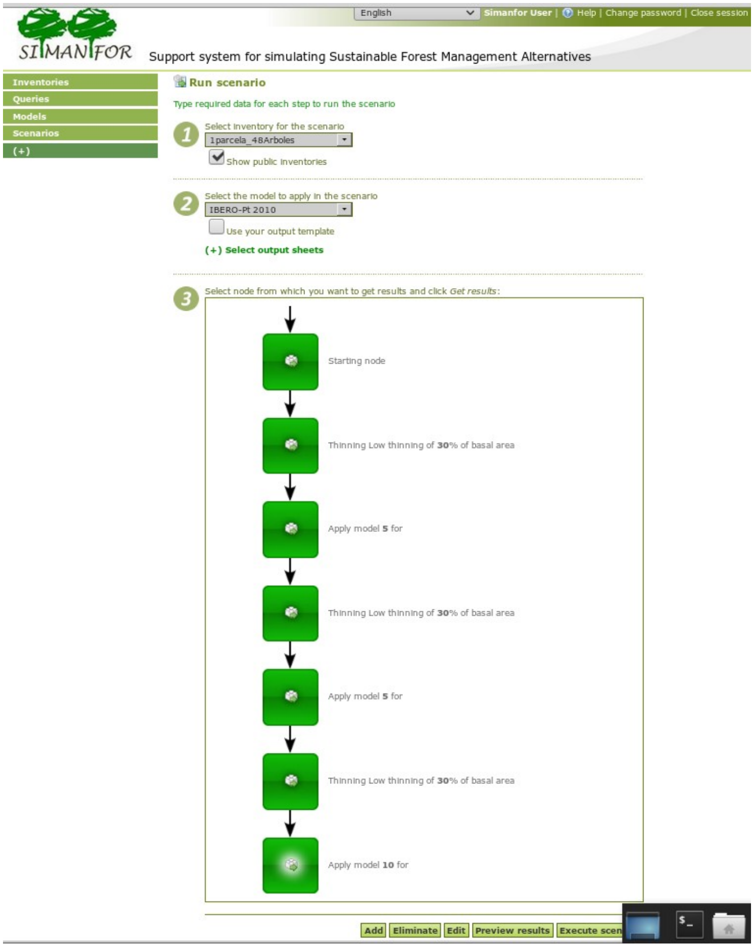

Figure 3 - High scenario for Simanfor

2.1.8 Classification of wood end use and quality

Based on taper equations, which can allow us to calculate the different possible cut for each bole, with optimum selection starting with bigest sizes.

The different end uses proposed for Cambric Pilot are the following, from small to large pieces of raw material needed:

- fuelwood, 8 to 20 cm in diameter and 1 m length,
- roundwood:
  - timber purlins, rafters, stakes and sticks, 8 to 15 cm in diameter and 2.5 m length,
  - Post or polewood, 10 to 45 cm in diameter and 6 m length,
- sawnwood, 20 to 200 cm in diameter and 1.2 m length,
- Veener, 15 to 160 cm in diameter and 2.6 m length,

Biomass estimation, total and for each component will be also provided. The equations developed by Ruiz-Peinado et Al. (2012 and 2013) will be used to achieve this result.

2.1.9 Data output format

Output data will be of four types:

1. a table with the same information as input data but one by each node performed by the simulation;
2. a summary of the simulation with the starting point, the interval of years simulated, and the silvicultural operations performed (thinnings)
3. a summary report with the same information as included in a forest yield table
4. a production table for end uses previously proposed

Outputs type 2, 3 and 4 will be also included in the ontology is being developed in the project by GSIC/EMIC and will be accessible as LOD in the Cross-Forest Endpoint.

2.2 FRAME: Forest fiRes Advanced ModElization

2.2.1 Overall Perspective

Goal: xxxxxxxxxxxxxxxxxxxx

Table 4 – FRAME pilot overall perspective

<table>
<tr><td>DESCRIPTION</td><td colspan="2">Use of an HPC-adapted propagation model for predicting forest fire behavior and spreading depending on variables related to weather (wind, temperature and humidity), vegetation and terrain, using data from Portugal and Spain with high spatial and time resolution for a cross-border approach.</td></tr>
<tr><td>PRECONDITION</td><td colspan="2">The propagation model is running in an HPC environment.
The databases of the Endpoint repository are accessible and published.
The external databases are published/available.
Access to the Endpoint repository and external databases is available.</td></tr>
<tr><td rowspan="6">NORMAL SEQUENCE</td><td>STEP</td><td>ACTION</td></tr>
<tr><td>1</td><td>Choose the ignition point, either by:
- typing X,Y coordinates
- drawing the point on a GIS interface (EMERCARTO).</td></tr>
<tr><td>2</td><td>Enter a range of different weather conditions: wind, temperature and humidity. These conditions can be real or simulated.</td></tr>
<tr><td>3</td><td>Run the HPC-adapted propagation model.</td></tr>
<tr><td>4</td><td>Modify weather/vegetation conditions</td></tr>
<tr><td>5</td><td>Run the HPC-adapted propagation model again and get a new iteration.</td></tr>
<tr><td>POSTCONDITION</td><td colspan="2">Quick mapping (and numerical information related) of fire behavior projections for the range of different weather conditions provided.
Weather/vegetation conditions can be modified and new iterations are obtained.
Geographical scope of fire behavior projections comprehends the Iberian Peninsula (Portugal and Spain).
Publication of results as Linked Open Datasets.</td></tr>
</table>

Figure 4 - Visual result of a propagation model

2.2.2 Outputs and user Interfaces

Fire spread and fire behavior models will be integrated to cover the calculation of overall fire spreading, fire spreading over particular elements (e.g. isolated trees, and bushes increasing the fuel load of a stand), crown fire initiation and spread (CFIS). Information from weather data sets, terrain, and outputs from other models (fuel, weather) income the integrated model, which returns fire spread variables (rate of spreading, fire intensity, flame length, residence time, etc.)

The expected output of FRAME pilot is to present a series of maps showing fire behavior projections for the range of different weather conditions provided, with a geographical scope that comprehends the Iberian Peninsula (Portugal and Spain).

These results will provide a useful management, prevention and training tool for Public Administrations and Emergency Professionals.

In this pilot, the resulting maps will be required to visualize in EMERCARTO, a GIS platform developed by Tragsa Group and employed for real-time emergency services monitoring, logistics and planning. EMERCARTO is employed by Tragsa Group and by several Public Administrations in Spain, and is accessible either from a PC or mobile devices. Users have different profiles and permissions.

2.2.3 Input data

a) From Cross-Forest LOD endpoint

Propagation models rely on information about fuel models. Information coming from Forest Inventories and Forest Maps, which will be accessible as LOD in the Cross-Forest Endpoint, will be the reference for building adapted fuel models that are efficient for running high resolution and simultaneous simulations over vast territories.

Forest variables such as Crown Bulk Density (CBD), and bush load and weight (UCO classification) will be obtained from

b) From external sources

To reach the main objective of FRAME, information about weather conditions and terrain with high spatial and time resolution will be used. The gathered data will allow the propagation model to run and to map fire behavior predictions based on real data. So, it is indispensable that we identify the external databases, and the information we will need to use as input data.

To carry out the predictions and to present information in EMERCARTO, it will be necessary to have access to datasets of some public administration organizations (Portuguese and Spanish), which have databases about weather and updated DTM. In Spain, AEMET and IGN will be key institutions as data providers. In Portugal, DGT and IPMA will provide the required data, as can be seen in the following table.

Table 5 – External datasets to be considered as input data for FRAME pilot

Themes	Variables	Data Sources	
		Spain	Portugal
Weather	Wind intensity and direction	AEMET WindNinja models	IPMA WindNinja models
	Temperature	AEMET	IPMA
	Humidity	AEMET	IPMA
Topography	Digital Terrain Model – DTM (x, y, z)	IGN	DGT

2.2.4 Data transformation

According to the data specificities and their providers, the project team will define the best workflow process to access, employ and update the data if necessary.

This project will provide various tools to answer several questions, as can be seen in the following Table.

Table 6 – FRAME pilot overview of data transformation required

Questions	File Required / Database required	Processing Rules
How will a fire behave if it starts now, at these coordinates, under current weather conditions?		
How would the fire behave under different weather conditions?		
How would the fire behave if vegetation conditions are modified (cutting, burning…)		
How would the fire behave if certain changes were		

performed in fuel conditions/amount/spatial distribution?		

References

Dunker, P., Barreiro, S., Hengevels, G., Lind, T., Mason, W., Ambrozy, S., & Spiecker, H. (2012). Classification of forest management approaches: A new conceptual framework and its applicability to European forestry. *Ecology and Society*, 17(4), 51.

Riofrío , Bravo F, Rio M (2019) Adapting the tree-level growth and yield model IBERO to mixed species stands (manuscript)

Ruiz-Peinado R, Rio M, Montero G (2011) New models for estimating the carbon sink capacity of Spanish softwood species. Forest Syst 20:176–188. doi: 10.5424/fs/2011201-11643

Ruiz-Peinado R, Montero G, Rio M (2012) Biomass models to estimate carbon stocks for hardwood tree species. Forest Syst 21:42–52. doi: 10.5424/fs/2112211-02193

